BIG DATA ANALYSIS OF COUNSELING CASES FOR YOUTH AT RISK OF SUICIDE

Seung Kwon[1], Mi-Jin Yang[2]*, Ju-Hyoung Yoo[3], Lae-Sun Kim[4]

[1]Korea Youth Counseling & Welfare Institute, Republic of Korea, ks1120@kyci.or.kr
[2]Korea Youth Counseling & Welfare Institute, Republic of Korea, kybj0910@gmail.com
[3]Korea Youth Counseling & Welfare Institute, Republic of Korea, loveyou@kyci.or.kr
[4]Korea Youth Counseling & Welfare Institute, Republic of Korea, iandq@kyci.or.kr

*Corresponding Author

Abstract

A close examination of the information promotion strategies of advanced countries, including South Korea, indicates the value of data, its analysis and usage. Big data analytics, in particular, attracts great attention as a method to fulfill personal and social demands by discerning a pattern for common behaviors and predicting the future. Along with drastic socioeconomic changes in South Korea, the suicide rate has increased dramatically and become the leading cause of death for those in their teens and twenties. This study investigates the features and contents of youth suicide by comparing social network data (69,886 cases in total) and suicide counseling cases from CYS-Net (1,463,737 cases in total). The results show that the youth in South Korea use the internet most at the beginning of the week, around the summer vacation, and from after school until night time. The main reasons for youth suicide are solitude and loneliness, worries about academic achievement or school admission, and school violence. More females disclose their problems to others than their male counterparts. 16 to 18 year-olds look for help the most from counselors for their problems. When it comes to location, youths in big cities, such as Gyeonggi-do, Seoul, Incheon, Daejeon, and Busan city, post more about suicide online. As these results show, youth suicide is a serious matter; thus, it is urgent to establish a youth suicide prevention system. This should be organized by population, age, and location, as well as related policy based on findings from this big data analysis.

Keywords: Big data, Suicide counseling, Social data, CYS-Net

1. INTRODUCTION

Along with drastic socioeconomic changes in South Korea, the suicide rate has increased dramatically and become the leading cause of death for those in their teens and twenties. Youth suicide exhibits some discrepancies compared to that of other age groups; it is a) difficult to predict as it occurs impulsively without a plan, b) caused by external factors rather than internal ones, and c) contagious as it leads to copycat or collective suicides (recited in Ju-mi Bae, Seung-yeon Lee, Eun-young Kim, 2010).

It is reported that suicide resulted from contagion, such as copycat suicide, mostly occur during the adolescent period or early adulthood (National Alliance on Mental Illness New Hampshire, 2006). Therefore, suicide is not just a problem for individuals but a problem for the whole society.

In Korea, suicide is ranked fourth in the cause of death statistics of Korea in 2013. It is twice higher than the average suicide rate of OECD member nations. Furthermore, youth suicide has intensified greatly and its rate has soared tremendously for the past few years. A report from 2013 states that the adolescent suicide rate in Korea is ranked number two among those of the developed countries, and this rate has increased by 57% over the past ten years (Hankookilbo, 2013.9.11).

Adolescence is a period when a level of anxiety and psychological vulnerability reaches its peak due to rapid physical and psychological changes. It brings about various problems, including identity crisis. For this reason, youth suicide shows different features compared to adult suicide. A major cause for adult suicide is a mental illness such as depression or schizophrenia, whereas in the case of youth, heavy stress or impulse to avoid difficulties provokes suicide even for those who function well. Also the amount of time from planning to an actual suicide attempt is shorter and impulsive. In addition, teenagers decide to take their own lives as a tool to manipulate, avenge, or ask for help (Min-kyung Hyun, Jong-min Woo, 2012). Youth's high level of anxiety and psychological vulnerability is linked to academic stress and career, relationship, and family

ISBN: 978-605-64453-2-3

issues, which heighten the possibility of committing suicide.
From 2010 to 2012, suicide is continually the number one cause of death in teenagers, and 11.2% of adolescents feel suicidal according to Statistics Korea. Moreover, the older the higher the rate of suicide as 10.7 ten to fourteen-year-olds and 25.8 fifteen to nineteen-year-olds out of 100,000 people took their own lives (Eun-kyung Yeo, 2012).
As youth suicide becomes a serious social issue, it has been researched in various studies. However, these studies mostly examine possible causal factors, such as depression, life stress and communication with parents. Studies dealing with a suicide prevention system or policy are insufficient (Eun-jin Jang, 2011).
Contrary to ongoing research, an effort to prevent and exterminate suicide at the government level is very concrete. The Ministry of Gender Equality and Family of Korea, responsible for youth policy, developed a network, called CYS-Net(Community Youth Safety-net) connecting youth-related organizations and facilities across the country. The Ministry of Education created and runs the 'Wee project' since 2008 for prevention, early discovery and professional intervention for suicidal youth. The Ministry for Health, Welfare and Family Affairs introduced 'the Act for the Prevention of Suicide and the Creation of Culture of Respect for Life' and administers the child and youth mental health promotion centers, health care facilities, local child care centers, and Dream Start for youth under age 18. Besides, police and fire departments affiliated to the Ministry of Security and Public Administration conduct a duty of emergency rescue for youths attempting suicide. Despite the previously mentioned endeavors of the government, the rate of youth suicide is still increasing and there is no solution available to ultimately solve this problem. In order to develop a practical suicide prevention policy, there is not only a need for in-depth research on suicidal youth but also a need to investigate a trend of the entire youth population. A detailed strategy formulated using knowledge of the whole trend would be the best way of creating an effective policy.
A close examination of the information promotion strategies of advanced countries, including South Korea, indicates the value of data, its analysis and usage. Big data analytics, in particular, attract great attention as a method to fulfill personal and social demands by discovering a pattern for common behaviors and predicting the future. In 2013, the Korean government announced the 'Government 3.0 Standard Plan Initiative' and support of data-based scientific policy development under the mission of smart governance utilizing big data. It is expected that big data analytics on youth suicide can detect the signs and patterns, and also devise an active model for copycat suicide prevention both online and offline. Therefore, this study investigates counseling cases for suicidal youth and attempts to create a plan for early suicide intervention by inspecting social data and youth counseling data.

2. METHODS

Data collected for the analysis includes 69,886 counts of social data derived from news, blogs, online cafe, SNS, online bulletin boards, and 30,398 counts of youth counseling data from CYS-Net including face-to-face counseling, a hotline, and cyber counseling. To gather the social data, the word 'suicide' was set as a search word and results assumed to be related to youth were extracted. The words 'president, own goal (which shares the same pronunciation with the word suicide in Korean), Werther, and movie' were put as stopwords to filter out unrelated data including articles on the suicide of public figures. For the youth counseling data, personal information was discarded. Those that contain the word 'suicide' and words that co-occurred with it were analyzed. Details about the analysis are included in Figure 1.

Figure 1. Sources of social and youth counseling data analysis

	Social Data Analysis	**Youth Counseling Data Analysis**
Date	**2012.1.1.~2012.10.18**	**2003.1.1~2013.11.30**
Buzz count	**69,886**	**30,398**
Channel	◦**News:** 214 online news web sites ◦**Blog:** Nate, Naver, Egloos, Daum, T-Story, Yahoo blog ◦ **Cafe:** Naver, Daum, Ppomppu, Card-gorilla, SLR Club ◦ **SNS:** Twitter, Me2day ◦ **Online bulletin boards:** Naver Kin, Ask Nate, Daum Tip, Nate Talk, Nate Pan, Daum Agora, Daum Miznet, Daum Tellzone, This is game, Human Rights Movement Sarangbang	◦ **CYS-Net** - Face-to-face counseling data - Hotline data - Cyber counseling data ◦**Youth Cyber Counseling Center** - Cyber counseling data - Cyber chat data

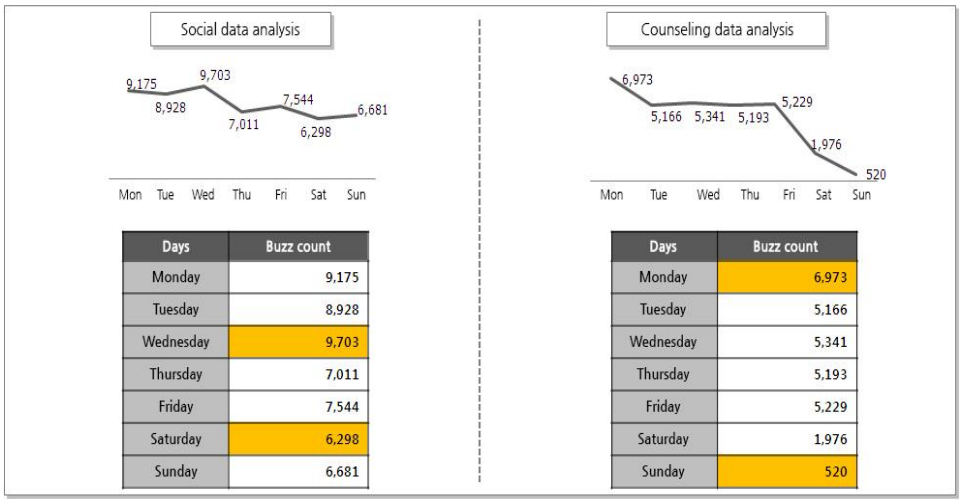
ISBN: 978-605-64453-2-3

3. RESULTS

3.1. Period when youth mention 'suicide' the most

The days of the week when 'suicide' is mentioned the most by youth are Monday, Tuesday and Wednesday in social data. It is only Monday in youth counseling data (figure 2). The word suicide is also the buzzword between 22:00 to 24:00 in social data and 16:00 to 18:00 in youth counseling data as demonstrated in figure 3.

Figure 2. Analysis of social and youth counseling data on a weekly basis



Days	Buzz count
Monday	9,175
Tuesday	8,928
Wednesday	9,703
Thursday	7,011
Friday	7,544
Saturday	6,298
Sunday	6,681

Days	Buzz count
Monday	6,973
Tuesday	5,166
Wednesday	5,341
Thursday	5,193
Friday	5,229
Saturday	1,976
Sunday	520

Figure 3. Analysis of social and youth counseling data on an hourly basis

Social data analysis

Hours	Buzz count
00시~02시	3,309
02시~04시	1,304
04시~06시	598
06시~08시	648
08시~10시	1,484
10시~12시	2,762
12시~14시	2,678
14시~16시	2,957
16시~18시	3,395
18시~20시	3,241
20시~22시	3,657
22시~24시	4,314

Counseling data analysis

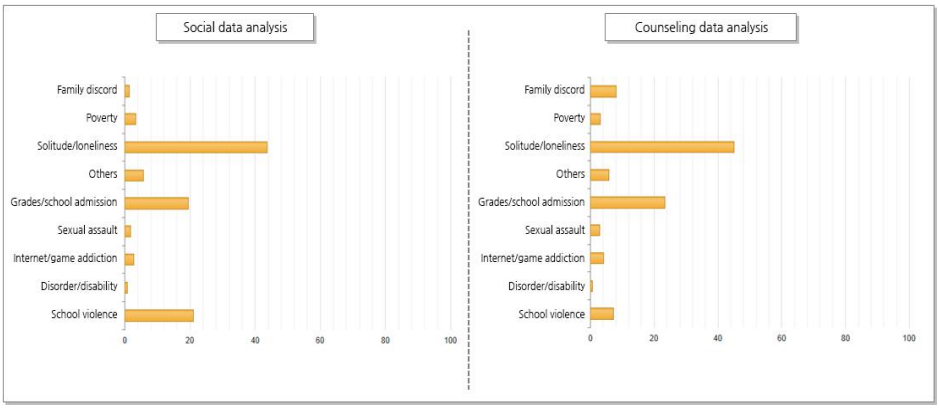
Hours	Buzz count
00시~02시	146
02시~04시	72
04시~06시	28
06시~08시	22
08시~10시	3,090
10시~12시	4,639
12시~14시	3,490
14시~16시	6,967
16시~18시	8,026
18시~20시	2,981
20시~22시	640
22시~24시	176

3.2. Keyword analysis of causes of youth suicide using social and youth counseling data

Figure 4 describes the analysis of keywords appearing with suicide in social and youth counseling data. According to the analysis, 'solitude/loneliness' co-occurred most frequently with suicide, followed by grades/school admission and school violence.

ISBN: 978-605-64453-2-3

Figure 4. Keyword analysis of causes of youth suicide using social and youth counseling data



In addition, 59.3% of females and 40.7% of males asked for help from counselors through CYS-Net. Middle school students held the highest rate with 47.5% while high school students accounted for 44.8% followed by 7.7% of elementary school students. In terms of location, youths in big cities, such as Seoul, Incheon, Daejeon, and Busan, posted more about suicide online. Details are indicated in figure 5 and 6.

Figure 5. Keyword analysis by gender and age

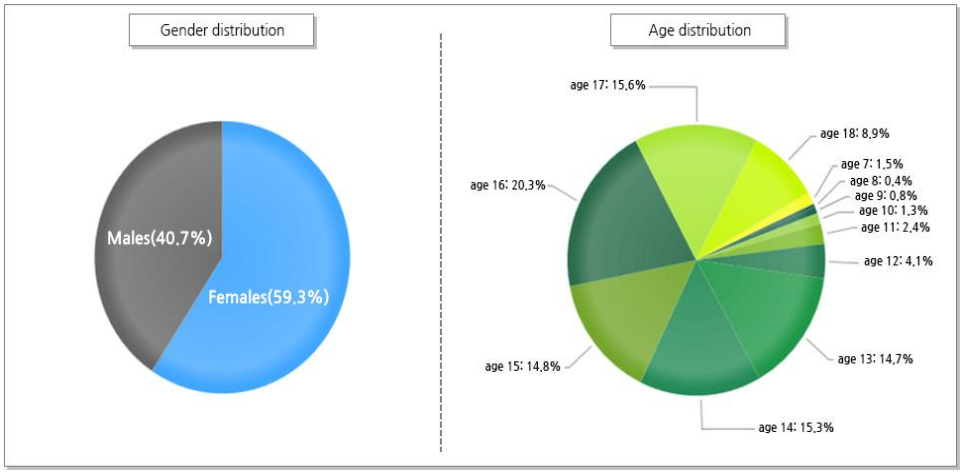
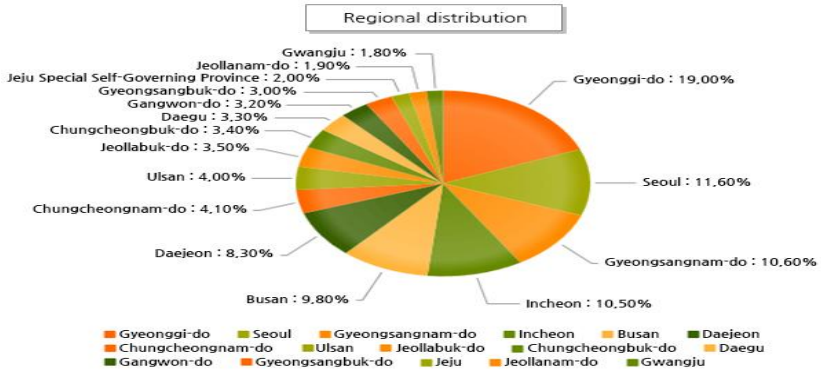


Figure 6. Keyword analysis by region

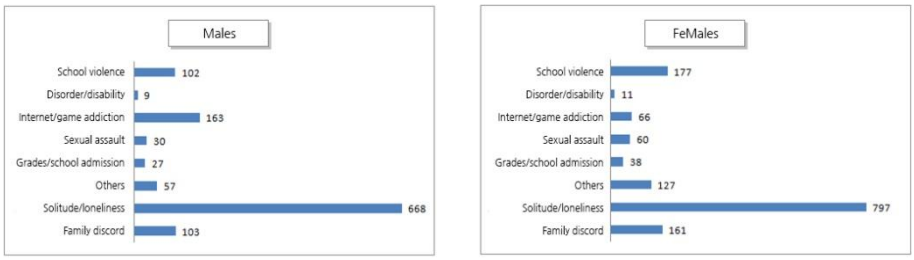


3.3. Keyword analysis of causes of youth suicide based on contents of counseling

Using youth counseling data containing face-to-face counseling, a hotline, and cyber counseling, a trend of keywords appearing with suicide was classified by age range from elementary school to college. Figure 8 shows that in the results derived from elementary students, solitude/loneliness co-occurred the most with suicide. It was followed by school violence and family discord.

Figure 7. Causal analysis using keywords appeared with suicide in youth counseling data (elementary school students)

ISBN: 978-605-64453-2-3



Rankings were the same for the middle and high school population which are included in figures 8 and 9 consecutively.

Figure 8. Causal analysis using keywords appeared with suicide in youth counseling data (middle school students)

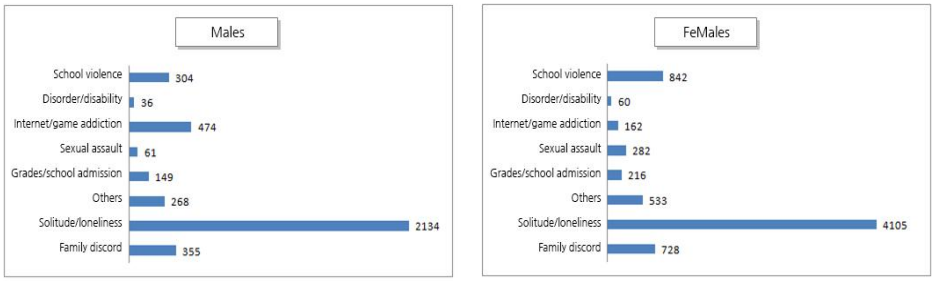
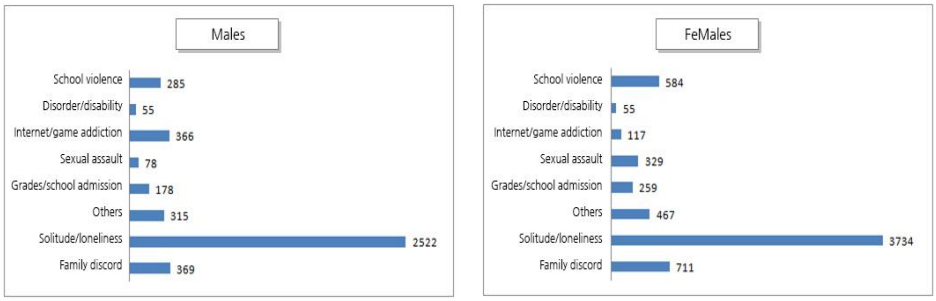
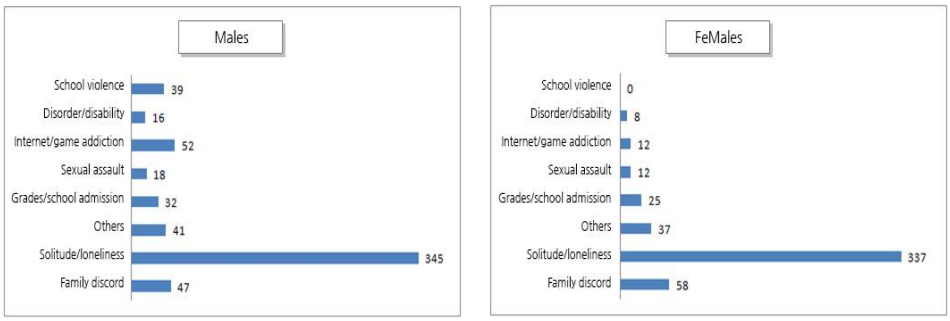


Figure 9. Causal analysis using keywords appeared with suicide in youth counseling data (high school students)



Finally, figure 10 demonstrates the results for college students; solitude/loneliness co-occurred the most with suicide and family discord coming next.

Figure 10. Causal analysis using keywords appeared with suicide in youth counseling data (college students)



4. DISCUSSION

The results of social and youth counseling data analysis show the necessity of establishing a systematic and flexible youth suicide prevention system which appropriately reflects the environmental factors. In this regard, this research has come up with several points about youth suicide. First of all, suicide is making the loudest buzz in the beginning of the week (Monday to Wednesday), and from after school until night time. Therefore, counseling channels for youth at risk of suicide, such as 'Helpcall 1388' or 'Cyber counseling center(www.1388.cyber.kr),' should be actively advertised at these times. An existing system also needs to

ISBN: 978-605-64453-2-3

be adjusted to serve at-risk-youth more readily by diversifying available counseling channels.
Secondly, age and gender are also critical factors to consider. The results of this study show that females disclose problems related to suicide more than their male counterparts. The rate of mentioning suicide increased from the second graders of middle school to those of high school. As youth enter middle school, relationship and academic issues generate severe stress. At this moment in time, there is no solution or a reliable network that can improve this situation. Moreover, even though it is a rather low 7.7% of the elementary school population that talks about suicide, it is also important to support this population in order to prevent proliferation of suicide with age. An age-specific suicide prevention system is needed so different age groups can be effectively served. For example, using a nation-wide counseling service that runs 24 hours, like 'cyber counseling center(www.1388.cyber.kr),' an intervention focused on characteristics by age needs to be prepared in advance. Additionally, a regional difference should also be considered when providing programs for training the professionals who actually intervene in person with an immediate response.
Third of all, in the keyword analysis to find the causes for youth suicides, 'solitude/loneliness' was ranked the highest as a buzzword in all youth groups from elementary to college population. Thus, the foremost thing to prevent youth suicide is to supply enough human resources who can play the role of a gatekeeper so those who are lonely do not make a radical choice. In this regard, not only professional psychotherapists and school counselors but also peer counselors, who can support youth at a close, immediate distance should be fostered. The keyword analysis also showed that issues generated from relationships, such as family discord or school violence, create great stress. To best support the youth at risk of suicide, counseling channels that can help them over 24 hours, like a cyber counseling center or local youth counseling and welfare centers, should be actively advertised to the youth population.
Lastly, it is imperative to have a government and any other youth-related facilities to be involved in preventing youth suicide. One such way would be to create a system that automatically forwards texts about suicide, violence or runaways on the youths' smart phones to their parents so their problems are known about early enough for immediate intervention. In addition, the system should be connected with on and offline counseling channels such as regional youth counseling and welfare centers, 'Helpcall 1388,' and cyber counseling centers as best practice.

5. SUGGESTIONS

This study intended to discover the trend of youth suicide by investigating social and counseling data that contain the word 'suicide' disregarding the context; thus, it is difficult to draw causal inferences for youth suicide from this study. The outcome of the study also cannot be generalized as the causal effects of suicide were inferred only by analyzing keywords that co-occurred with suicide on data. However, this study effectively does utilize trends drawn from the big data analysis and suggests a practical and competent suicide prevention and early intervention plan. In order to be a useful database for policy making, big data related to youth crisis should be accumulated continuously in order to be utilized productively.

REFERENCE

Bae, J., Lee, S., & Kim, E. (2010). Development of the Post-intervention Manual on Adolescent Suicide. *Youth Counseling Research from Korea Youth Counseling and Welfare Institute,* vol.2010.

Yeo, E. (2012). *A Study on Theories and Researches on Youth Suicide. The Study of Child-Family Therapy,* vol. 11.

Jang, E. (2011). Critical Review on the Current Study of Adolescent Suicide in Korea. *Journal of Future Adolescents,* 8(4).

Hyun, M., & Woo, J. (2012). A Study on the Regional Risk Factors for Youth Suicide. *Seoul: National Evidence-based Healthcare Collaborating Agency.*

Hankookilbo. (2013, September 11).

National Alliance on Mental Illness New Hampshire. (2006). *Frameworks Youth Suicide Prevention Project: Postvention.*

ISBN: 978-605-64453-2-3